Coding of Missing Data (UOE source):

The correct use of the missing codes is an *essential* condition to ensure the integrity of the data. Different reasons why there are no data in a particular instance need to be distinguished in the statistical analyses and the reports and must therefore be distinguished in the data tables.

Each cell for which there is no valid data value must be assigned one of the following five missing codes:

- category not applicable (**a**);
- data included in other categories (**x**, **xr**:..., **xc**:..., **xa**:...);
- data not available (**m**);
- data value nil (**n**);
- data value negligible (**n**...).

✗ Note that for reasons inherent in EXCEL$^©$ the missing codes had to be changed from earlier INES data collections.

Category not applicable (a)

If a certain category or cell in the tables generically does not apply to the educational system(s) in your country then assign the code "**a**" to all cells referring to this category (or cross-classification of categories). This implies that data for these categories data do not even hypothetically exist.

☞ However, if a category applies but there are no data available or if the data for this category are included in another category then do *not* use this code.

✎ *For example, if in a country there are no independent private schools, then for all variables the cells referring to independent private schools should be assigned the code "a". If, on the other hand, independent private schools do exist but data are not available for these, then one of the other missing codes must be used. For example, if the data for independent private schools are included in the data referring to the total of all private schools then the sub-categories referring independent private schools should be coded to "included in another category" instead.*

You should provide an annotation to categories which are not applicable to your education system (↗ see page 3-**Error! Bookmark not defined.**).

Data included in other categories (x...)

Sometimes certain data are not available (even though the corresponding categories apply) but the data are included in other categories of the tables. Most

frequently this occurs in situations where data are not available for certain sub-categories and can be provided for sub-totals or totals only.

The correct analytic treatment of data that are included in other categories is not possible without the knowledge in which category these data are included. The target category may thereby be in another row of the table, in another column of the table, or in another row *and* column. In order to process such instances in a computerised way, the target category must be identified as described below in the corresponding missing code:

Data included in sub-total (x)

If data are not available for certain sub-categories but the data for these sub-categories are included in the corresponding nearest row- and column totals of the data collection tables (but not in any other sub-categories which are not sub-totals) then enter the code "**x**" into all of the corresponding cells.

✎ *For example, central government expenditures on student loans (row-code C11 in table FINANCE1) may not be available separately for the, primary, lower secondary, and upper secondary level of education (UOE column-codes,2,3, and 4) but may be available for the total of these levels (column-code 5). In such a case the code "included in nearest sub-total (x) would be assigned to the cells with the UOE row- and column-codes in C11/2, C11/3, and C11/4.*

However, there are other situations where data are not available for a sub-category but these data are included in another sub-category other than a sub-total or total.

✎ *For example, the data on government expenditures on student loans may not be available separately for the pre-primary and primary level of education but only for the total of both (for which table FINANCE1 does not offer a separate sub-total).*

In such cases you should use one of the following four codes:

Data included in another row (xr:...)

Use the code **xr:** *row_code* if the data for this category are included in another row of the same column but not in a cell of any other column (other than a column total). *row_code* thereby refers to the UOE row-code of the target row which is indicated in each table (and which is usually different from the internal EXCEL® sequential row-number).

✎ *For example, if at the upper secondary level (column-code 4 in table FINANCE1) the expenditures on student loans from local sources (UOE row-code L11) are not available separately from the expenditures on student loans from regional sources (row-code R11) you may report the total of local and regional expenditures in the row referring to regional sources (row-code R11) and insert the code "xr:L11" in the row referring to local sources (row-code L11).*

☞ For any such cases it is assumed that the data value is also included in the row- and column totals which belong to the target category. This must therefore not be separately indicated.

✎ *In the preceding example, one would assume that the column referring to the total financial aid to students from regional sources (row-code R12) would also contain the expenditures on student loans from local sources (row-code L11) which were included in row R11.*

☞ Whenever the data of two rows need to be merged, choose the row for which you expect the larger numeric value as the target row.

☝ If a category is not applicable, then do not use any of the "x" codes but use the code "a" instead.

Data included in another column (xc:...)   Use the code **xc:** *column_code* if the data for this category are included in another column of the same row but not in a cell of any other row (other than a row total). ***Column_code*** thereby refers to the UOE column-code of the target column which is indicated in each table (and which is usually different from the internal EXCEL® sequential column-number).

✎ *For example, a country may not be able to distinguish between expenditures on student loans (UOE row-codes C11, L11, R11, G11 in table FINANCE1) at the lower secondary level (column-code 3) and the upper secondary level (column-code 4) but still may have data for the total of both levels. In such a case you may report the combined expenditures for the upper secondary level and code the cells C11/3, L11/3, R11/3, and G11/3 to xc:4 respectively.*

For any such cases it is assumed that the data value is also included in the row- and column totals which belong to the target category. This must therefore not be separately indicated.

☞ Whenever the data of two columns needs to be merged, choose the column for which you expect the larger numeric value as the target column.

☝ If a category is not applicable, then do not use any of the "x" codes but use the code "a" instead.

*Data included in another or multiple cells (**xa:...**)*	Use the code **xa:** *row_code* **:c** *column_code*: ***annotation*** if the data for a category are included in another cell (which does not belong to the same row or column) or if the data are included in another category but none of the above situations applies (✎ *for example, because the data are included in multiple rows and/or columns*). In this case you need to supply a separate textual annotation which indicates in which category or categories the data are included.

✎ *For example, in table FINANCE1 for local government expenditures there may be no data available at the primary level of education (UOE column-code 2) for direct expenditures for independent private institutions (UOE row-code L3). However, if these data are included in the total of expenditures over all levels (row-code L11, column-code 10) but not in the total of expenditures for pre-primary, primary, and secondary, then assign the code xa:L11C10 to the cell L3/2 and provide a corresponding annotation.*

☞ If a category is not applicable, then do not use any of the "x" codes but use the code "a" instead.

*Data not available (**m**)*	Sometimes data are generally not available for certain classification categories and are not included in any other cells of the table (even though these data could, in principle, be collected). In such cases you should assign the code "m" to all cells referring to the corresponding classification categories.

☞ However, if the data for these classification categories are included in other categories of the tables or in the corresponding row- and column totals, then do not use the code "m" but use one of the "x"-codes instead.

*Magnitude nil (**n**)*	If a data value is nil then assign the code "n" to the corresponding cell.

☞ If a category is not applicable and therefore the data value would be zero for a cell, then do not use the code "n" but use the code "a" for "category not applicable" instead.

☞ Do not use the value 0 (zero) to indicate that a value is nil.

*Magnitude negligible (**n**...)*	In other cases data may not be available but the magnitude may be known to be negligible with respect to the values in the other categories of this variable. Such instances should be assigned the code "n<magnitude>" followed by an indication of the

estimated magnitude. ✎ *For example, if the value of a category is estimated to be smaller than 120, then the code would be "n120".* This will be important when such data are used for international comparisons because, obviously, a value which is negligible in the United States may not be negligible in Luxembourg.

✎ *For example, some countries have very small numbers of independent private schools for which they collect no data but they may, for example, know that there are less than 120 teachers employed in such schools as compared to 325000 teachers in government dependent private and public schools. In such case the code n120 would be assigned to the categories of educational personnel in independent private schools.*